Benchmarking inappropriate empirical antibiotic treatment

G. Kariv[1,2], M. Paul[2,3], V. Shani[1,2], E. Muchtar[1,2] and L. Leibovici[1,2]

1) *Department of Medicine E, Rabin Medical Centre, Beilinson Hospital, Petah-Tikva, Israel* 2) *Sackler Faculty of Medicine, Tel-Aviv University, Petah-Tikva, Israel and* 3) *Unit of Infectious Diseases, Rabin Medical Centre, Beilinson Hospital, Petah-Tikva, Israel*

Abstract

Inappropriate empirical antibiotic treatment for severe infections is associated with increased mortality. Superfluous treatment is associated with resistance induction. We aimed to define acceptable rates of inappropriate empirical antibiotic treatment. We included all prospective cohort studies published between 1975 and 2009 reporting the proportion of appropriate and inappropriate empirical antibiotic treatment of microbiologically documented infections. Studies were identified in PubMed and in reference lists of included studies. Funnel plots were drawn using the proportion of inappropriate empirical treatment as the effect size. A pooled estimate of inappropriate empirical antibiotic treatment was calculated using a β-binomial model. Control limits were calculated with the overdispersion factor technique and 20% winsorized data. Heterogeneity was assessed through subgroup analysis for categorical moderators and meta-regression for continuous variables. Eighty-seven studies, comprising 92 study groups, with 27 628 patients met inclusion criteria. The pooled rate of inappropriate empirical antibiotic treatment was 28.6% (95% CI 25.4–31.8). Funnel plot analysis yielded a dispersed graph with only 37 (40%) studies falling within the control limits. Using the overdispersion factor technique with 20% winsorizing, 79 (86%) studies fell within the control limits. None of the clinical or methodological factors could explain the large heterogeneity observed. The funnel plot presented can be used to benchmark rates of inappropriate empirical antibiotic treatment. Based on the control limits found, at least 500 patients should be evaluated before establishing a local rate. Lower and higher than expected rates might indicate overly aggressive treatment or poor performance, respectively.

Keywords: Benchmark, control charts, funnel plot, inappropriate empirical antibiotic treatment
Original Submission: 28 January 2012; **Revised Submission:** 31 May 2012; **Accepted:** 6 June 2012
Editor: D. Raoult
Article published online: 18 June 2012
Clin Microbiol Infect 2013; **19:** 629–633
10.1111/j.1469-0691.2012.03965.x

Corresponding author: L. Leibovici, Department of Medicine E. Rabin Medical Center, Beilinson Hospital, Petah-Tiqva 49100, Israel
E-mail: leibovic@post.tau.ac.il

Introduction

Inappropriate antibiotic usage affects both the single patient and the community. From the single patient perspective, inappropriate empirical antibiotic treatment is associated with a significant increase in mortality [1]. From the community perspective, superfluous antibiotic use leads to economic cost and, more importantly, an ecological problem of resistant bacteria induction [2]. Therefore, empirical antibiotic treatment of suspected moderate to severe bacterial infection does not aim at 100% coverage of all possible pathogens, but is an attempt to strike a balance between coverage and the ecological impact of broad-spectrum antibiotics [3]. We aimed to examine whether this balance is uniform in different practices.

We reviewed studies reporting on the rate of inappropriate empirical antibiotic treatment for documented bacterial infections, following a predefined protocol. We performed a meta-analysis of inappropriate empirical treatment rates. Our aim was to define an acceptable range of rates based on currently reported rates and to assess whether there are factors that underlie a justified deviation from the acceptable range. These results can be used to benchmark appropriateness of empirical antibiotic treatment.

Methods

Data sources

We searched PubMed for studies looking at the percentage of empirical antibiotic treatment. Prospective studies (defined as those where at least data collection was performed prospectively) published between 1975 and 2009 were included if addressing adults (≥18 years) with microbiologically documented infections and treatment was selected by clinicians. Studies where treatment was defined by study protocol or those addressing patients treated with specific antibiotics were excluded. The definition of appropriate empirical antibiotic treatment was treatment that was given before the results of the cultures were known and matched the *in vitro* susceptibility of the pathogen. We permitted the inclusion of studies where up to 10% of microbiologically documented infections cannot be tested *in vitro*; in these cases the study definitions for appropriateness were accepted. We excluded studies assessing meningitis, endocarditis or viral infections; and studies that recruited <50 patients or were published in languages other than English. We used the following search clause: ((antibiot* OR antimicrob* OR anti-bacter* OR antibacter*) AND (approp* OR inapprop* OR adequate OR inadequate)) AND ((cohort* OR prospect*) NOT retrospect*) NOT Review[ptyp] AND 'adult'[MeSH Terms]. References from identified studies were also scanned.

Data extraction

Two reviewers independently extracted the data from included studies. In case of any disagreement between the two reviewers, a third reviewer extracted the data and consensus was reached. We extracted data on appropriate empirical treatment (definition, timing of treatment, number of appropriate and/or inappropriate treated patients). In addition, we extracted data on study characteristics to allow the examination of factors affecting the rate of inappropriate empirical antibiotic treatment. We collected data on settings, study years, study objectives, follow-up duration, patient characteristics, pathogens and source of infection. In cases where data were published in multiple studies, the data were included only once.

Data synthesis and analysis

The pooled estimate of inappropriate empirical antibiotic treatment was calculated using various methods, including fixed and random effects models, simple binomial model and β-binomial model. Fixed and random effects models were performed using COMPREHENSIVE META ANALYSIS version 2.2, Simple binomial and β-binomial calculations were performed using SAS software and SAS BETABIN MACRO (http://www.qistats.co.uk/BetaBinomial.html). Funnel plots were drawn using the proportion of inappropriate empirical treatment as the effect size. The funnel plot graph uses five lines. The central horizontal line is the pooled proportion estimate. The selected model for the pooled proportion estimate was the β-binomial model. This model is proposed for combining overdispersed binomial data across multiple, heterogeneous studies [4]. The upper and lower lines present the control limits, calculated as 2 SD and 3 SD (~95% and ~99.8% prediction limits). Control limits are calculated using the overdispersion factor technique and 20% winsorized data [5]. Any point falling outside the control limits is an outlier, suggesting a special cause for the variation. The funnel plots were prepared based on a funnel plot EXCEL template downloaded from Easter Region Public Health Observatory (ERPHO, http://www.erpho.org.uk/).

Analysing the cause for variation of the effect size used subgroup analysis for categorical moderators and meta-regression for continuous variables (Comprehensive Meta-analysis Version 2, Biostat, Englewood, NJ, USA (2005)). Correlation analysis between variables was performed using SPSS (version PASW STATISTICS 17.0, Release 17.0.2, SPSS Inc., Chicago, IL, USA).

Results

The search for potentially eligible studies yielded 1053 references. Eighty-seven studies reporting rate of inappropriate empirical treatment and meeting the inclusion criteria were included (Fig. 1). These publications comprised 92 study

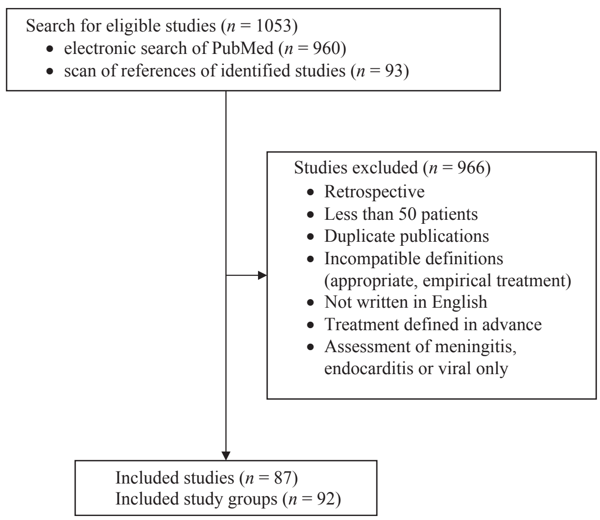

FIG. 1. Study flow.

TABLE 1. Pooled rates of inappropriate empirical antibiotic treatment

Method	Estimate (%)	Lower 95% CI	Upper 95% CI
Simple binomial	29.19	28.65	29.72
β-Binomial	28.65	25.45	31.85
Meta-analysis (fixed)	31.4	30.8	32
Meta-analysis (random)	26.3	23.7	29

TABLE 2. Comparison between presented funnel plots

	Within limits (%)	Below 2σ (%)	Above 2σ (%)	Below 3σ (%)	Above 3σ (%)
Basic funnel plot	37 (40)	12 (13)	6 (7)	18 (20)	19 (21)
Overdispersion, 20% winsorized	79 (86)	5 (5)	2 (2)	2 (2)	4 (4)

Data are presented as number of studies (%).

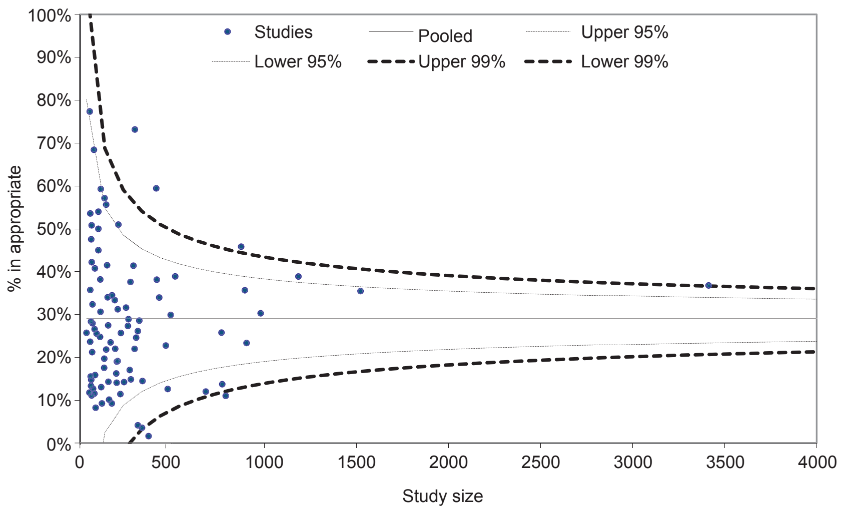

FIG. 2. Funnel plot using overdispersion factor, 20% winsorized. The plot presents the percentage of inappropriate empirical antibiotic treatment against the study size.

groups, which were used to develop the pooled estimate and the funnel plot. The studies are described in Appendix S1 and referenced separately (Appendix S2).

The studies included a total of 27 628 patients and were published between 1977 and 2007. The studies were carried out in 23 different countries, predominated by Spain with 34 studies and USA with 24 studies (only a few studies were multi-centred, Appendix S1). Forty-one studies (45%) were conducted in an intensive-care unit (ICU) and the remaining studies were carried out exclusively or predominantly in a non-ICU setting. Forty-six studies (50%) assessed bacteraemia, 24 (26%) assessed pneumonia and the remainder assessed a mix of microbiologically documented infections. Most (71, 77%) of the studies evaluated pathogens combined, but 15 (16%) and 6 (7%) focused on gram-negative and gram-positive infections, respectively.

The reported proportions of inappropriate empirical treatment varied from 1.61% to 77.36%. Significant heterogeneity between studies was observed (I^2 = 95.6%, p <0.001). The random effect meta-analysis yielded a pooled proportion of inappropriate treatment of 26.3% (95% CI 23.7–29). Results for the pooled proportion in the different models are presented in Table 1. The pooled proportion estimate used in the funnel plot is the β-binomial calculated average, 28.6% (95% CI 25.4–31.8). Data presented using a funnel plot yielded a dispersed graph with only 37 (40%) of the studies falling within the control limits. When using the overdispersion factor technique with winsorizing of 20%, a more suitable funnel plot was produced, leaving only a few studies as outliers (Table 2 and Fig. 2).

We assessed the effects of various study characteristics on results. These included demographic variables (country—divided into low, medium and high prevalence of resistant pathogens, age, ICU or non-ICU, place of acquisition), pathogen, type of infection and percentage of bacteraemia, study methods (design, primary aim and appropriate definition) and years of data collection (Appendix S1). The proportion of inappropriate empirical treatment was associated only with the rate of infections caused by gram-negative bacteria in the study with borderline statistical significance (increase of 1.006, 95% CI 1–1.012; for a 1% increase in gram-negative infections, p 0.05), mixed effect meta-regression. None of the other variables could explain the heterogeneity in reported rates of inappropriate empirical treatment.

Discussion

The objective of this study was to calculate and present, based on the available literature, a method for benchmarking inappropriate empirical antibiotic treatment. A pooled estimate rate of 28.65% inappropriately treated patients was calculated from 87 prospective studies, comprising 27 628 patients, with highly significant heterogeneity. Using the funnel plot technique, only 40% of studies fell within the 95% confidence limits of the pooled rate estimate. We could not explain the heterogeneity between rates in individual studies using a large dataset of the cohort characteristics. Specifically, the rate of inappropriate treatment did not change during the last decades; and was not dependent on the severity or type of infection (except for a marginally significant influence of the percentage of gram-negative infections). We therefore used the overdispersion factor technique to incorporate the heterogeneity in the benchmark funnel plot. Winsoring of 20% was necessary to obtain a funnel plot where 86% of studies fell within the 95% confidence limits of the pooled rate.

Quality control and performance measurement is gaining focus in healthcare. Measures are used to answer questions like 'How are we doing?', 'What should we be doing?' and 'Have the changes we have made led to improvement?' [6]. A funnel plot is a graphical tool mainly used for detecting bias in meta-analysis [7]. A funnel plot is a chart in which data of interest (the indicator/proportion) are plotted against a denominator representing the precision such as volume, sample size or size of the study cohort. In addition, control limit lines are drawn, creating the funnel. In recent years funnel plots have been proposed as one of the preferred tools for quality control and performance comparisons [8–15]. Their main advantages over other methods are that they do not promote ranking, they accommodate variable sample sizes, and the data are visually displayed for an intuitive interpretation. When presenting data in a funnel plot, common versus special cause variation is intuitively presented, whereas using a table to show the same data would promote 'best' and 'worst' values.

This paper presents a practical and reliable method for benchmarking the rate of inappropriate empirical antibiotic treatment. The method relies on a large amount of data being available and on robust statistical methods. There are several limitations to our analysis. We could not find reasons for the large heterogeneity. Causes can range from data quality issues to local resistance profile or quality of care. We analysed many variables extracted from the studies as possible moderators, but each was tested separately in univariate analysis. Perhaps a mix of several variables (e.g. study carried out in ICU in recent years) is needed to explain heterogeneity. Our inclusion criteria can be contested. We aimed to include all studies whose primary, prospectively defined, objective was to examine the rate or effect of inappropriate empirical antibiotic treatment. We probably missed studies reporting rates as part of other research objectives. The included studies are not necessarily representative: they were performed in hospitals with an interest in the topic and with the resources to conduct prospective studies and with an interest in publishing them. Many small studies reported a low rate of inappropriate treatment, and we cannot rule out publication bias.

Further research should address healthcare organization volume and superfluous antibiotic usage. We believe that the funnel plot generated here can be used as the start of a benchmarking effort by hospitals that aim to test their antibiotic practice. Use of ongoing databases might provide better data for benchmarking than the use of published data.

The funnel plot can also be used to test the adequacy of antibiotic treatment in trials of antibiotic and non-antibiotic treatment of patients with septicaemia. A small cohort will defeat the purpose of the funnel plot, as the confidence intervals are wide for small numbers. The 95% CI boundaries for a cohort of 500 patients are 15–42%, and we recommend using a cohort of at least 500 patients for benchmarking.

Transparency Declaration

All authors have no conflict of interest to declare.

Funding

This work was supported in part by a grant from the Israel Science Foundation No.1193/07.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Characteristics of included trials.

Appendix S2. References of included trials.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

1. Paul M, Shani V, Muchtar E, Kariv G, Robenshtok E, Leibovici L. Systematic review and meta-analysis of the efficacy of appropriate empiric antibiotic therapy for sepsis. *Antimicrob Agents Chemother* 2010; 54: 4851–4863.
2. Patrick DM, Hutchinson J. Antibiotic use and population ecology: how you can reduce your 'resistance footprint'. *CMAJ* 2009; 180: 416–421.

3. Leibovici L, Shraga I, Andreassen S. How do you choose antibiotic treatment? *BMJ* 1999; 318: 1614–1618.
4. Young-Xu Y, Chan KA. Pooling over dispersed binomial data to estimate event rate. *BMC Med Res Methodol* 2008; 8: 58.
5. Spiegelhalter DJ. Handling over-dispersion of performance indicators. *Qual Saf Health Care* 2005; 14: 347–351.
6. Siriwardena AN. Using quality improvement methods for evaluating health care. *Qual Prim Care* 2009; 17: 155–159.
7. Egger M, Smith GD, Schneider M *et al.* Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 1997; 315: 629–634.
8. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005; 24: 1185–1202.
9. Noyez L. Control charts, Cusum techniques and funnel plots. A review of methods for monitoring performance in healthcare. *Interact Cardiovasc Thorac Surg* 2009; 9: 494–499.
10. Wouters MW, Krijnen P, Le Cessie S *et al.* Volume- or outcome-based referral to improve quality of care for esophageal cancer surgery in The Netherlands. *J Surg Oncol* 2009; 99: 481–487.
11. Kunadian B, Dunning J, Roberts AP *et al.* Funnel plots for comparing performance of PCI performing hospitals and cardiologists: demonstration of utility using the New York hospital mortality data. *Catheter Cardiovasc Interv* 2009; 73: 589–594.
12. Young RJ, Holmes EM, Casson IF *et al.* The North West Diabetic Pregnancy Audit: a practical system for multi-centre diabetic pregnancy audit. *Diabet Med* 2008; 25: 496–500.
13. Davies E, Mak V, Ferguson J, Conaty S, Møller H. Using funnel plots to explore variation in cancer mortality across primary care trusts in South-East England. *J Public Health (Oxf)* 2008; 30: 305.
14. Hurley JC. Profound effect of study design factors on ventilator-associated pneumonia incidence of prevention studies: benchmarking the literature experience. *J Antimicrob Chemother* 2008; 61: 1154–1161.
15. Gale CP, Roberts AP, Batin PD *et al.* Funnel plots, performance variation and the Myocardial Infarction National Audit Project 2003–2004. *BMC Cardiovasc Disord* 2006; 6: 34.